DOCUMENT RESUME

ED 095 371 95 CE 001 935

AUTHOR	Harmon, Lenore W.
TITLE	Technical Aspects: Problems of Scale Development, Norms, Item Differences By Sex, and the Rate of Change in Occupational Group Characteristics.
INSTITUTION	Aries Corp., Minneapolis, Minn.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	74
CONTRACT	OEC-0-72-5240
NOTE	44p.; For related documents, see CE 001 926-934, Draft
EDRS PRICE	MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS	Career Choice; Equal Opportunities (Jobs); Interest Research; Interest Scales; *Interest Tests; Measurement Techniques; Occupational Choice; Sex Differences; *Sex Discrimination; Sex Stereotypes; Test Bias; *Test Construction; *Test Reliability; *Test Validity; Vocational Interests
IDENTIFIERS	Sex

ABSTRACT

The paper considers how the technical aspects of interest measurement can contribute to sex bias and suggests guidelines for evaluating interest inventories for sex bias and for developing interest inventories which minimize sex bias. An overview is given of interest measurement techniques: in selecting a pool of items, developing scales, and norming the scales. The issues raised at each step are surveyed. Interpretive practices are briefly discussed. Suggestions are offered for minimizing sex bias at each of the three major technical steps in developing a new interest inventory, and for its interpretation. Twelve specific guidelines are presented in summary, including these: The content of interest inventory items should not imply that any occupation or activity is more appropriate for one sex than for the other; scores on the same scales should be available for both men and women; norm groups for occupational scales should contain men and women in equal proportions or the effects of both groups should be statistically equalized; published interpretive materials should indicate clearly that sex norms for homogeneous scales do not imply anything about the predictive validity of the scales. (AJ)

This paper was funded under a contract to the ARIES Corporation (#OE-C-52-7240) from the National Institute of Education of the U. S. Department of Health, Education, and Welfare. The views expressed do not necessarily reflect those of the ARIES Corporation, the National Institute of Education, or any agency of the U. S. government.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Aspects: Problems of Scale Development, Norms, Item Differences By Sex, and the Rate of Change in Occupational Group Characteristics

DRAFT

BEST COPY AVAILABLE

Lenore W. Harmon
University of Wisconsin--Milwaukee

ED095371

The world of work is divided into two categories, "women's work" and "men's work" and anyone who does not know the difference can consult such widely divergent sources as children's primers (Women as Words and Images, 1972), the local want-ads, or United States Department of Labor statistics (1969). Individual differences between men and women have been cited as the reason for this dichotomy (Parsons, 1965) as have politics and economics (Bird, 1968; Millett, 1969; Firestone, 1970). Actually, there are few jobs which require skills and abilities which are possessed exclusively by one sex or the other. In fact, there are few such skills (Lewis, 1968). Thus, it follows that there are few instances in which it is appropriate for an individual "to limit...his or her consideration of career solely on the basis of gender."

Historically, the oldest interest inventory still in use, the Strong Vocational Interest Blank, was developed in the 1920's and 1930's on the assumption that there are two categories of work and two categories of workers. Reporting on disappointing early attempts to measure women's interests, Strong commented, "The writer's hunch is that it will be found that men can be better differentiated respecting a larger number of occupations than can women. The primary reason for this belief is that the interest technique necessitates a fairly homogeneous criterion group. At the present time far too many women enter an occupation as a stop-gap until marriage. Consequently, they take a job because it is convenient, not because they intend to continue in it indefinitely. The result is that most occupations contain a considerable number of women who would not be there if they had

CE 001 935

selected an occupational career as men do. Any sampling of such occupations gives a rather heterogeneous group of women. Occupational scales based on such criterion groups cannot be expected to differentiate very well." (Strong, 1943, p. 129).

Strong developed separate inventories for men and women. His assumptions probably reflected the convictions of his time, but they also influenced future researchers. The latest Mental Measurement Yearbook (Buras, 1972) showed that the male form of the Strong Vocational Interest Blank has been utilized in 1,099 studies since it was introduced in 1927, while the female form of the Strong Vocational Interest Blank had been utilized in 168 studies since it was introduced in 1933. About 15 times more research has been directed toward studying men's interests than women's interests, if we make the questionable assumption that each reported study represents an equal unit of research. Strong assumed that women's interests were difficult to measure and different from men's interests: he seemed to convince a majority of researchers and inventory developers. The interest inventories available today rest on many of the same assumptions that he utilized.

One purpose of this paper is to consider how the technical aspects of interest measurement can contribute to sex bias in that interest inventories may be used to limit consideration of career on the basis of sex. Another purpose is to suggest guidelines for evaluating interest inventories for sex bias and for developing interest inventories which minimize sex bias.

Overview of Interest Measurement Techniques

There are three major technical steps in developing any interest inventory and sex bias can enter at any point. First, a pool of items are selected; second, scales are developed; finally, the scales are normed.

Items

Most interest inventory items utilize occupational names and/or occupational activities as items. Names of school subjects, non-occupational activities, or personal characteristics are used as items to a lesser extent. If the items imply sexual stereotyping in occupations, i.e., "policeman" or "sales lady", they promote sexual bias both in the responses they elicit and in the image of the occupation they project. Strong (1962) suggested guidelines for selecting good items for interest inventories but his assumptions and his measurement technique did not lead him to include the elimination of sex bias.

Scale Construction

After a set of items have been developed, they are administered to a group of people. The responses of this group to the set of items provide the basic data from which interest inventory scales are constructed. There are basically two kinds of interest inventories and two ways to construct them. An understanding of the two types of scales and the way they are constructed is necessary to an understanding of the potential for sex bias in each type. Scales may be based on either external or internal criteria. The composition of the group which is used to provide data for scale buildin and the way that data is treated determines the type of scale which results.

Scales based on external criteria utilize the item responses of a group of people already employed in a specific occupation (the criterion group) to provide a comparison with either a reference group of people

employed in a large number of occupations or with the individual being tested. These scales are called occupational scales and named after the occupation represented in the criterion group. A review of all interest inventories listed by Buros in the two most recent editions of the Mental Measurements Yearbook (1965, 1972) revealed that only five of the 52 contain occupational scales based on external criterion groups. The five include the Kuder Preference Record - Occupational (Form D), the Kuder Occupational Interest Survey (Form DD), the Minnesota Vocational Interest Inventory, and the Strong Vocational Interest Blank for Men (Form T399), and the Strong Vocational Interest Blank for Women (Form TW398).

Since these inventories are widely used it is worthwhile to examine their scale building methods closely. The occupational scales of the Kuder (Form D), the Minnesota Vocational Interest Inventory and both forms of the Strong Vocational Interest Blank are all developed in a similar manner. For each item, the percentage of an occupational criterion group who make each response are compared with the percentage of a reference group who make each response. For example, on the Strong Vocational Interest Blank item, "Biologist", individuals may respond "Like", "Indifferent", or "Dislike". To determine whether that item should appear on a scale for female veterinarians, a comparison is made between the percentage of the occupational criterion group (female veterinarians) and the percentage of the reference group (Women-in-General) who make each of the three responses. Figure 1 illustrates how weights are assigned to each response for the female Veterinarian scale. Only items which differentiate between the two groups by a large percentage, usually 15 percent or more, are included.

The occupational scales of the Kuder Occupational Interest Survey (Form DD) also utilize criterion groups of people employed in an

occupation. However, the responses of the individual taking the inventory are compared to those of the criterion group and expressed as a type of correlation coefficient between the criterion group and the individual (Zytowski, 1973). No reference group is used.

Most interest scales based on external criterion utilize occupational groups of one sex or the other but not both. All interest scales which have been developed using external criteria rest on the assumption that sex is very important to building the scales.

Scales based on internal criteria, also called homogeneous scales, are developed by using some method of clustering the interest inventory items. The clustering method may reflect the subjective judgment of the test developer, the dimensions of a theoretical model, or the results of a factor analysis. The important difference from occupational scales is that the clustering process does not make use of an external criterion such as membership in an occupational group in assigning items to scales. The group of people tested to provide basic data about the inventory items are not specified by occupation although they may be specified by age or sex. Their responses are used to provide information on how the items cluster. Items which are highly correlated for the group tested are placed together in scales. The resulting scales are usually more limited in number than occupational scales based on an external criterion and they have names which are more general ("Mechanical", "Clerical", and "Social", as opposed to "Engineer," "Secretary", and "Social Worker".)

Another important difference between the two types of scales is that most homogeneous scales based on internal criteria are not sex specific as are the occupational scales. These internally referenced scales which

are designed to reflect general dimensions of interest present in a general population, usually make no assumptions about sex and bias is not introduced in the scale building process. Bias might be introduced, however, if the group used to develop item statistics for the clustering procedure were atypical in some way, i.e., contained only one sex or the other. Generally, bias is not introduced in this way and more potential for bias in inventories developed using internal criteria exists in scale norming than in scale building.

Norming Scales

Occupational scales based on external criterion groups of people employed in the occupation usually utilize the occupational criterion group as a norm or comparison group. The score of an individual reflects some degree of similarity or dissimilarity to men or women in the occupation when an occupational scale is employed. Thus, a degree of concurrent validity is built into the externally referenced scales. The assumption that occupational scales require separate criterion groups by sex leads to scales which require separate norming by sex if the scales are to make any interpretive sense.

The norm group for internally based scales is usually the same type of general population group utilized in scale building. The age range of the norm group is the same as that of potential test takers. The scales have implied or demonstrated content and construct validity but there is neither concurrent nor predictive validity associated with the comparison or norm group. The homogeneous scales of inventories of this kind present a sort of occupational typology (which may or may not be carefully documented in content and construct statistics). They show individuals their type(s) compared with others who are similar in age and/or sex. It has seemed reasonable to present separate norm groups by sex on this kind of interest inventory because the distributions

of scores for men and women are usually different. However, it also is reasonable to ask whether the use of separate norm groups by sex introduces sex bias.

Issues Raised

This review of the techniques of interest measurement raises many issues about potential sex bias in interest measurement. First, it is important to to establish whether there are sex differences in response to items because item responses are the building blocks of interest measurement. If there are no sex differences at the item level, none of the interest measures which are built upon those items need to take account of sex. If there are sex differences at the item level it is important to determine whether they are related to either item content or item format, and to determine whether separate pools of items are needed for men and for women.

Another related issue is whether occupational scales need to be developed for each sex separately. Why are many homogeneous scales developed without reference to the sex of the respondents?

It follows that we need to know whether either occupational or general norm groups should be separated by sex.

The goals of interest measurement are to help individuals explore their interests in comparison with others and to promote good life planning. If any of the practices which seem to imply sex bias in interest measurement are not actually necessary to these goals, then sex bias does exist in interest inventories.

Issues Related to Items

Sex Difference in Item Response

Campbell (in press), has been able to use the large bank of interest data at the Center for Interest Measurement Research to show that there are

large sex differences in response to some items, such as "operating machinery," "decorating a room with flowers," and "repair electrical wiring." These differences appeared between men and women from the early adolescent years to the adult years; between groups of adult men and women in the same occupations, and between groups of men and women in the same occupations in 1930 and in 1968. Sex differences in item response seem stable over age and time and they do not disappear when occupation is held constant. Johannson and Harmon (1972) examined responses to all the items common to the male and female forms of the Strong Vocational Interest Blank with occupation held constant and found that an average of 42 percent of the items showed large sex differences. Within occupations, the percentage of items showing large sex differences varied from 30 to 58. The percentage of items exhibiting large sex differences when the Men-in-General and Women-in-General groups were compared was 44. These data were based on recently tested occupational criterion groups and they indicate a substantial amount of difference in the way men and women respond to typical interest inventory items.

Separate Item Pools

Since there are documented sex differences in response to interest inventory items, the practice of providing separate sets of items for each sex should be examined. A recent review of Buros Mental Measurements Yearbooks for 1965 and 1972 showed that eight inventories or sixteen percent of all the inventories listed in those years have separate forms for males and females. Not all of them were inspected but it seems clear that the only reason to provide separate forms is that the item pool differs for men and women. The underlying assumption seems to be that men and women play different occupational roles which cannot be described by one set of interest

inventory items.

Where separate item pools exist critics have cited the presence of items descriptive of high level technical and business activities in the item pools for men and the presence of items descriptive of domestic and clerical activities in the item pools for women. The Strong Vocational Interest Blank (Form T399 and Form TW398) have over half their items in common yet one of the issues that led to the development of the Strong Campbell Interest Inventory, to be published in 1974, was the separate item pools. In a memo to the Strong Vocational Interest Blank Advisory Board regarding the need for an immediate revision of the Strong Blanks, combining them into one form, Campbell wrote in January of 1972, "The SVIB does tend to perpetuate stereotypic roles for men or women, at the expense of women, both by the kinds of items included, and the kind of information provided in the profile."

Since many reputable interest inventories have been able to develop effective scales without using separate item pools, it can be concluded that separate item pools are a source of sex bias in interest measurement in that they call attention to sex differences unnecessarily.

One of the important side effects of the separate item pools on a well researched inventory like the SVIB is to preclude easy comparisons between the item responses of men and women. Even though half the items are identical, they are numbered differently in the two forms and appear on different scales, creating nearly prohibitive difficulties for anyone wanting to compare male and female responses. This author believes that systems are designed to accommodate projected needs.

The fact that male-female comparisons across the Strong Blanks are so difficult suggests that no one in over 40 years considered them important until now.

Identical items pools for men and women are essential to avoid introducing sexual bias and to facilitate research on sex differences in item responses. Where sex differences in item response are found, it is necessary to explore their real impact on interest inventory scales before one draws conclusions about the effect of sex differences in measuring interests.

Item Content

Interest item content can imply the "correct" sexual identity of the respondent who endorses it. Few men would dare say they like the occupation "saleslady." Certainly items which include the words "man," "woman," "girl," "boy," or "lady," are sexually biased. Other items which do not explicitly involve a sexual stereotype nevertheless may elicit stereotyped images and responses because of cultural bias. For example, the item "nurse" is probably perceived as a feminine occupation and responded to as a feminine occupation although the name of the occupation does not imply an "appropriate" sex.

There has been some controversy over the best type of item content for interest inventories. Strong (1943) built his inventories with several types of content such as occupational names, activities, school subjects and personal characteristics. Holland (1965) developed an inventory, The Vocational Preference Inventory, with items which are all occupational titles. He argued (1973) that occupational names stimulate reliable vocational stereotypes in the test taker. Campbell, (in press) found that occupational titles as items separated occupational groups more effectively than items on the Strong Vocational Interest

Blank which utilized other types of content. Kuder (1970a) argues persuasively against the use of occupational titles.

Now we need to determine if sexual stereotyping of occupations is more closely related to one type of content than another. For instance, if an occupational name like "Physician" is perceived as masculine, while an activity like "watch open heart surgery" is not, using items based on occupational activities may be a subtle way of combating the culturally induced sexual stereotyping which each individual brings to the testing situation.

Item Format

There has been considerable controversy about item format or the way the item is presented (Kuder, 1970a, Dolliver, 1968; Campbell, in press). The most popular formats include presenting one item and asking the respondent to endorse it or not and presenting two or three items and asking the respondent to choose between them or rank them. Both formats have been accused of reducing validity. If men and women respond to item formats differently, item format can contribute to bias. Strong(1943) presented evidence of minimal differences in response bias between men and women, but it is not clear whether they responded to the same items. The issue is complex and of lesser importance than some of the others, but it might be investigated.

Issues Related to Scale Development

The task in developing scales for an interest inventory is to provide scores which will reflect the interests of an individual in a way that is useful in career planning. Obviously scales which contain sexual bias cannot adequately fulfill this function. Given the fact

that scales are built from items and items do differentiate between the sexes, it is important to explore the impact of sex differences on the scales.

Scales Based on External Criteria

For interest inventories utilizing external criteria in scale development, the crucial issue is what the criteria should be. Usually, the criteria is membership in some occupational group--by sex. There are two subissues, however. First, there is the question of whether it is useful to employ separate criterion groups of each sex where members of both sexes are readily available. Second, there is the question of what to do when members of only one sex are available for a criterion group.

Separate sex scales. The belief that men and women in the same occupation do different things has been expressed by Johnson(1970). Intuitively his assessment of the situation seems correct; male social workers become administrators and teachers, while female social workers see clients; male lawyers practice corporate law while female lawyers practice social law. We do not know whether these arrangements of people within occupations are political or whether they are based on real sex differences in interests.

Both the Strong Vocational Interest Blank and the Kuder Occupational Interest Survey scales are based on separate criterion groups of males and females (even if it requires two scales for one occupation). The implication is that sex differences are very important in occupations and in measuring interests. In the study (Johansson and Harmon, 1972) where large numbers of items were found to differentiate between men and women in the same occupation, a further analysis was performed.

The investigators found that a relatively small proportion of the items which differentiated between the sexes in an occupation actually appeared on the male or female Strong Vocational Interest Blank scale for that occupation. Thus, the effect of sex differences in item response is not as great as it appears from the number of such items because they do not all appear on occupational scales. Johansson and Harmon concluded that it might be possible to build good scales by combining males and females in an occupation and comparing their item responses with those of a combined in-general or reference group. Since the Kuder Occupational Interest Survey uses all the items on each scale (Zytowski, 1973), male-female item differences might have a greater effect on scales if criterion groups of people in the same occupation were combined.

Combined sex scales. Some preliminary work on building occupational scales based on both sexes is being done by the author and Pat Weber of the Center for Interest Measurement Research. The new Strong Campbell Interest Inventory has a common item pool so the effect of male-female item differences on scales can be studied directly. Three tentative scales were developed; a scale utilizing the comparison between male veterinarians and a male reference group, a scale utilizing comparisons between female veterinarians and a female reference group, and a scale utilizing the comparison between a criterion group of male and female veterinarians weighted equally and a reference group of males and females weighted equally. Table 1 lists the items from the Occupations section (items 1-131) which appeared on each of the three tentative scales. Eight items appeared on all three scales, twenty items appeared on two scales (all twenty on the scale for combined sexes

and ten each on the male and female scales), five items appeared on the male scale only, seven appeared on the female scale only, and one appeared on the combined scale only. These findings demonstrate that for veterinarians it is possible to develop an occupational scale for combined sexes which has an adequate number of items and which will correlate with both the single sex scales. Appropriate validity studies have yet to be computed. Whether or not other occupational scales can be derived in the same way is not clear. Strong (1943) studied this problem by scoring men and women on scales for both sexes. He concluded that the men's and women's scales could be combined in some cases but not in others. Interestingly enough, he did not try it. More research is still needed. If valid scales can be built using combined criterion and reference groups, it will not mean that the interests of males and females in an occupation are the same, but that sex is not an important enough part of the criterion to require a separate scale.

Weighting sexes in criterion groups. Another kind of problem arises in building occupational scales where there are only a few members of one sex employed in the occupation and/or available to be in the criterion group. The procedure used in scale building should minimize the effect of male-female differences in item response unless it can be demonstrated that they are importantin building an effective scale. Although the idea of composing each criterion group of the same proportion of males and females as are found in the occupation makes some sense at first, further reflection shows that it is only a modification of the approach already in use. If 90 percent of an occupation is composed of males and that proportion is maintained in the criterion group we would expect the scale to be

more representative of men's interest than of women's interests.

Actually, it is not even that simple where the Strong Vocational Interest Blank is involved. Since the method employed literally removes the interest common to one sex or the other from consideration (Dolliver, 1968), atypical interests for males may be typical interests for females and result in findings like those of Stanfiel (1970) who showed women obtained more high scores on the men's form than the women's form of the Strong Vocational Interest Blank. Campbell (in press) presented data which showed that males score higher on many scales developed for females than on the scales developed for males in the same occupation while females scored higher on many scales developed for males than on the scales developed for females in the same occupation. It might be possible to apply a sort of correction factor if the differences were consistent across all occupations. Unfortunately, there is considerable variation in the amount of difference between men and women in scores on male and female occupational scales.

We can conclude that the best way to minimize sex bias is to weight males and females equally in criterion groups and reference groups even if that is not the proportion which reflects the status of the occupation. Schlossburg and Goodman (1972) have indicated that there are enough people of the "opposite" sex in many occupations to provide adequate criterion groups of males and females in atypical occupations. Criterion groups are usually selected by finding some concentrated source of people committed to their occupation, such as a professional organization, union, or licensing board. Scarcity of members of one sex in an occupation dictates a more extensive and more expensive search for criterion group members. While the number of individuals needed in a criterion group varies in a way that is not predictable beforehand,

200 individuals is probably a reasonable minimum (Harmon, 1968). To provide a good criterion group with both sexes respresented equally it is probably necessary to have a minimum of 200 individuals of each sex. If this minimum is met, the item response percentage for each sex should be averaged.

One sex scales. This procedure will not work where less than 200 persons of each sex are available, so the question of whether to build scales on a single sex (and how to use the many which have already been built) arises. It seems unreasonable to build more scales based on one sex when they may contain sex bias. After equal opportunity has been on operation for a few decades we may be able to determine if there are occupations which only attract members of one sex. Until then, our efforts can be expended in building occupational scales based on criterion groups of both sexes.

Where scales have already been developed using criterion groups of one sex, we should probably continue to use them for the appropriate sex until they can be converted to scales for both sexes. There has been considerable controversy about whether to use scales developed one one sex with the other. Darley and Hagenah (1955, p. 71) suggested using the men's form of the Strong Vocational Interest Blank in conjunction with the women's form for women who have a "high degree of career motivation, maturity, and ability..." Strong, (1943, p. 576) was convinced that "it is much better to score a sex on its own scales." The Association for Measurement and Evaluation in Guidance Commission on Sex Bias in Measurement (1973) took the position that using "wrong"

sex scales is psychometrically meaningless. The data from Stanfiel (1970) and Campbell (in press) suggest they are right.

Kuder (1970, b) showed that there are respectable correlations (medians from .74-.81) between women's scores on thirteen sets of Occupational Interest Survey scales for men and women in the same occupation. Although the level of their scores on the male and female scales might "differ considerably", "scales that yield the highest scores in one set generally yield high scores in the other." Thus, it appears that one can use "wrong" sex scales on the Kuder Occupational Interest Survey if one explores the highest scores on those "wrong" sex scales comparing their content but not their level with that of the highest scores on the correct sex scales. This procedure cannot be generalized to inventories with scales which were not developed as the Kuder Occupational Interest Survey scales were developed without additional study. It would seem more fruitful to exert the same effort in developing new scales where scales based on one sex are the only scales available.

Note that there is seldom a problem over testing men on occupational scales developed on criterion groups of women. The scales which are available for women but not for men do not represent the most prestigious or highly paid occupations in our society, which may explain why men do not need to use them. Alternatively, the cultural prohibitions against men being interested in "women's" occupations may be much stronger than cultural prohibitions against women being interested in "men's" occupations.

Criterion group stability over time. Since collecting criterion groups is expensive, it is reasonable to ask whether, once collected, scales based on them can be appropriately used over long periods of time. Campbell (1971, Chapter 9) presented four different kinds of evidence that the people in various occupations do not differ much over 30 years in their measured interests. Only one of the studies involved women subjects. However, the evidence strong-

which do not change much over time. This finding probably applies equally to men and women.

General changes in the popularity of Strong Vocational Interest Blank items over time have been documented for women by Hill and Campbell (1969) and for men by Campbell (1968). Since interests of those in occupations are more stable, these findings apply more directly to the composition of reference or in-general groups. If reference groups are to be used the problem of cultural change in item response must be investigated further. Unfortunately, studies of differential rates and types of cultural change between men and women have not been done as

t'ey require longitudinal date and a common item pool for men and women. Apparently, researchers on the Kuder inventories are in a position to provide this information earlier than any other researchers.

The foregoing discussion leaves unresolved the question of what to do if there is only an occupational scale for the "wrong" sex or no scale at all available where one is needed in counseling. The availability of occupational scales can be a source of sexual bias, in that one can be discouraged from considering a career or a whole set of careers by the absence of an appropriate interest inventory scale. The only solution to this problem in both the long and short range comes from the use of homogeneous scales based on internal criteria, since occupational scale building will always lag behind changes which occur in the world of work.

Scales Based on Internal Criteria

Homogeneous scales appear on the Strong Vocational Interest Blank and the Minnesota Vocational Interest Inventory, which also have scales based on external criteria, as well as on dozens of interest inventories have homogeneous scales exclusively. The set of homogeneous scales avail on a given inventory suggest a model of the organization of vocational interests.

Many researchers have addressed the question of the factorial struct of interests. Two of the earliest studies, by Thurstone (1931) and Criss and Daniel (1939), found somewhat different factors for men and women. Both studies found "science", "language", and "people" factors but Thurst found a "business" factor for men while Crissey and Daniel found an "inte in male association" factor for women. Scales which contributed to the l ter included: housewife, office worker, stenographer, and nurse.

More recently, factor studies have concentrated primarily on the structure of men's interests as reviewed by Super and Critis (1962). So have attempts to understand the conflicting information which sometimes arises from the use of two inventories (King, Norrill and Powers, 1963; Zytowski, 1968, 1969, 1972; Kuder, 1969; O'Shea and Harrington, 1971, O'Shea Lynch and Harrington, 1972).

The most promising data on the structure of interests is that of Holland (Holland, Whitney, California and Richards, 1969.) and Cole (1973; Cole and Hanson, 1971; Cole, Whitney and Holland, 1971; Cole and Cole, 1970) in which a structure is proposed, a methodology established, and applications of the structure to occupations, to men's interests and to women's interests are substantiated. The results show that there is an interest structure common to men and women which corresponds to the circular two-dimensional arrangement of interests proposed by Roi (1956) and Holland, et al (1969). Both inventories based on internal criteria and inventories based on external criteria were included in Cole's analyses.

Cole (1973) suggested an extremely valuable method of surmounting the problems of sex bias inherent in using sets of externally referenced scales which have no scales for occupations women do not traditionally enter. It is to use the available scales from these inventories to locate the individual's interests in the circular structure and to generalize from that position to the full range of occupations which occupy nearby positions in the structure. Homogeneous scales on the Vocational Preference Inventory (Holland, ~~1965~~ 1965a), ACT's Vocational Interest Profile (1972), and the new Theme scales on the Strong Campbell Interest Inventory (Campbell, in press) correspond directly to the structure and can be used to locate individuals' interests in the structure.

The process does rest on an important assumption which should be investigated. Since the occupational locations in the structure are based on the vocational preferences of young people, not employed adults, the method does not possess any demonstrated predictive validity. The work of Campbell and Holland (1972) and Hansen and Johansson (1972) established concurrent validity for scales related to Holland's model (1966, 1973) which corresponds to Cole's structure by showing that they separate occupational groups. Like most homogeneous scales, they do not differentiate as well between occupational criterion groups as occupational scales. However, there is enough data regarding the construct validity of Holland's formulation (Holland, 1973, Chapter 5) to suggest that more evidence of predictive validity will be forthcoming. At the present time, Cole's method for using interest inventories appears to be the best way to procede, but predictive validity must be established.

In general, homogeneous scales have not been shown to have much predictive validity. Zytowski (in press) reviewed studies which have attempted to establish predictive validity for homogeneous scales. He conducted a study of the Kuder Preference Record utilizing two methods. His conclusions were that the Kuder (Form B) was predictive of future employment but not accurate enough to use in individual counseling. His subjects were men but his study illustrates the point that the type of instrument which has least potential for sex bias because of common item pools and common scales for both sexes has the least predictive validity (Harmon, 1973).

Issues Related to Norms

Norms for interest scales with external criteria are usually implied by that criteria. If the criterion group is made up of persons of one sex so is the norm group. The same issues about the appropriateness of this procedure arise in discussing both norms and scale construction. The inappropriateness of scoring individuals on scales developed on the "wrong" sex was discussed earlier. The only way to insure that the norms on interest inventories based on external criteria are not sexually biased is to make sure that the scales are not sexually biased.

Norms for homogeneous scales based on internal criteria are usually developed by scoring a group of people within some age range on the scales. Most homogeneous scales such as those on the Vocational Preference Inventory, the Kuder (Form C), the Ohio Vocational Interest Survey, and the Strong Vocational Interest Blanks have separate norms for men and women. While norms do not confer validity, they are used to compare individuals with a group which will provide a meaningful comparison. In the case of interest inventories these comparisons are used as a basis for discussion about future plans. D'Costa (1972) and Goldman (1972 (a)) have advocated that interest inventory results be used as aids in discussion and vocational exploration. However, as soon as they are used predictively, the norms become very important. We know that it is psychometrically defensiveable to provide separate norms when the scores for two groups (for instance, male and female) differ appreciably. However, the practical meaning of sex-segregated norms is obscure. For instance, if a high school senior woman scores high on a homogeneous scale for Machine Work compared with a high school

senior women how should she interpret this information? One might guess that she would score lower on norms for high school senior men but on the Ohio Vocational Interest Survey report form (where a Machine Work scale appears) the male norms are not available to her. Should she be encouraged to consider machine work because she scores high compared to her same sex peers, encouraged not to consider machine work because she probably scores lower than her opposite sex peers, or encouraged to consider her interest in machine work compared to people who do and like machine work. In the latter case, the norms offer her no information at all.

The problem is that general sex norms for homogeneous scales confer no criterion related validity but we tend to use them as though they do. It would probably be better to report scores on homogeneous scales simply as a code type, as in Holland's Self Directed Search (1965, b), encouraging individuals to explore their strongest interests and avoid comparisons with other people completely. If results of homogeneous scales are to be used predictively the scores should be validated. Concurrent validity could be established by establishing norm groups of employed workers in various occupations as with externally developed scales. However, the question of how the sexes should be represented in such norm groups is unresolved. They could be represented equally, proportionately, or separately, and only further study will show how it should be done. It is certain to be an expensive and difficult task.

Although Cole (1973) developed her structure of women's interests utilizing scales which were developed and/or normed for women only, the author feels strongly that it would be more fruitful to explore the predictive validity of her suggested procedure than to collect more occupational norm groups for homogeneous scales.

This type of research should be possible both at American College Testing Program and Center for Interest Measurement Research. Both have a large mature bank of data on individuals, in the sense that the data are old enough to support meaningful longitudinal studies. The eventual occupations of students who took the Vocational Interest Profile at ACT can be ascertained. At the Center for Interest Measurement Research the task is even easier, since older Strong Vocational Interest Blanks for people whose eventual occupations are known can be scored with the scales based on Holland's theory (Campbell and Holland, 1972; Hansen and Johansson, 1972).

If the use of homogeneous scales normed for women and the occupational structure presented by Cole are predictively valid they should be used. The use of code types to locate an individual in the occupational structure without further reference to norms should also be explored. If these procedures do not prove to be predictively valid, we will have to norm homogeneous scales on employed groups in an attempt to establish concurrent validity.

While a discussion of validity seems out of place in a discussion of norms, it is clear that norms on interest inventories have no purpose at all if they are unrelated to validity.

The discussion of norms brings up the question of how age is relate to criterion groups and norm groups. Campbell (1971, Chapter 5) has demonstrated that there are a number of Strong Vocational Interest Blank items which differentiate between men of various ages. He used them to develop an Age Related scale. He was also able to identify scales where scores increased or decreased with age in a large fairly representative sample of men. Apparently, even though individual Strong Vocational Interest Blank scores are quite stable over time (Campbell, 1966, 1969) there are group differences in item response and scale scores over time.

Since women, at least at this time, probably utilize vocational counseling on a different schedule than men, the effects of these differences may be important. Astin and Myint (1971) and Harmon (1970) have shown that career women are difficult to distinguish from non-career women until the woman is an adult, after the time when the information is needed for career planning. Many women do seek career counseling as adults with grown children, rather than as late adolescents.

It is appropriate to ask whether interest inventories provide adequate measures of interest for "returning" women. The question of appropriate norm groups and occupational groups for older women is crucial.

Given the data presented by Campbell it would appear to be best to have norm groups or criterion groups composed of a wide range of ages. This is usually the case in inventories based on external criteria since occupational status is more important than age. On inventories based on internal criteria age as well as sex norms are often used. Few of them have adequate adult norms, high school and college norms being much more typical. Homogeneous scales can be used with returning women to establish an individualized order of preference among vocational dimensions without reference to norms. This preference may be more useful than comparisons which are potentially effected by age differences. These issues should be considered in developing interest scales to be used with women because of the different life patterns they are likely to follow.

Interpretive Practices

These are obviously complex issues which it is important for the counselor to understand. Some of the problems can be alleviated by choosing the most appropriate inventory for each individual. If the client is a 14-year-old girl the problem is quite different from that of a 50-year-old woman, and it would seldom be appropriate to select the same interest inventory for both. In general, it is important to select an inventory with an unbiased item pool. The counselor should

not reinforce sexual stereotyping by presenting the client with test materials which appear to divide the world of work into two classes--male and female. The report of scores or interest profile should not contribute to sexual stereotyping either, but report forms for most of the major inventories are currently unable to meet this criterion. They either contain separate scales for each sex, or separate norms for each sex. Since we are not sure at this time whether separate norms and separate scales are necessary, since there almost certainly are complex influences at work to make the use of "wrong" sex scales and norms questionable, how to present interest inventory scores is a problem. It seems silly for psychologists to have to admit to clients, that our professional eyesight has been afflicted by a form of double imagery, (seeing the world of work as a sexual dichotomy) which has resulted in a set of psychological instruments which we now realize (being on our way to better eyesight) are largely uninterpretable. The best proposal before us seems to be to use the instruments available to assign women clients to a position in Cole's (1973) structural model (which approximates Holland, Whitney, Cole and Richards' theoretical hexagon). This procedure will require that the practicing counselor keep a close eye on the research literature of vocational interest measurement. If some evidence of the predictive validity of this approach does not appear within the next five years, it should be abandoned.

Two Psychometric Fantasies

Suppose that the origins of interest measurement lead back to a group of men and women who were not operating under the assumption that some jobs are men's jobs and other jobs are women's jobs. They would have devised a pool of items to differentiate between people in various occupations and people in a general reference group or to differentiate people who had high interest in one interest factor or cluster from people who had high interest in another. Certainly, this idea could have occurred in the 1920's when the early feminists had scored a major victory and earned the vote. If sex had been an important variable in measuring interests in various occupations or interest factors, the researchers would have found out early. If they started with the assumption that all work is available to all people, they would not have relegated the problem of how to treat real sex differences to unimportance. Instead, they would have worked to insure that sex differences entered the measurement process only where they were really appropriate and had something to contribute toward helping individuals learn about their interests and how those interests relate to jobs. But this fantasy is too idealistic, one cannot change the past.

One can influence the future, so it may be more productive to fantisize in that direction. If I were given an unlimited grant today to build a new interest inventory which would minimize sex bias, how would I procede?

Items

Obviously, I would avoid items which imply an "appropriate" sex. I would attempt to determine whether items containing occupational names or occupational activities elicit more responses based on sexual

stereotypes. Data is being analyzed now which may help answer this question. Conroe (1974) asked 50 male and 50 female college students to respond to the Strong Campbell Interest Inventory in a "typically masculine" and "typically feminine" manner. Since the Strong Campbell Interest Inventory contains both types of item content, it will be possible to study the responses for each condition to compare which class (names or activities) elicits the largest differences between them. Actually, the data will allow for comparisons between all the types of content on the Strong Campbell Interest Inventory.

Since I know that some items are more highly endorsed by men than by women, I would want to balance the number of each type in my inventory. (Association for Measurement and Evaluation in Guidance Commission, 1973). This would avoid the appearance of bias toward one sex or the other and contribute to face validity for both sexes. It would also insure that a preponderance of items more highly endorsed by one sex or the other on any scale, could not be attributed to the base rates or number of such items available in the item pool.

Scales

Since there is no clear answer as to whether scales based on external or internal criteria are best, I would build both types. First, I would administer my items to a large randomly selected group of adults and identify interest factors or clusters. I would expect that factors similar to the types proposed by Holland (1966) would appear. It is clear from Holland's review (1973) that many researchers are finding that both interest and occupational data fit his model.

Then I would build some occupational scales. I would attempt to select occupations for study which; 1) are least likely to imply sexual

seterotyping when their names appear on a profile or report form, 2) are representative of major areas of interest, and 3) are representative of major levels of occupations. Diamond (1971) has shown that sex differences are more important at low occupational levels than at higher occupational levels. I would avoid occupational groups with extreme splits in sexual composition. If a group had less than 20% of one sex or the other I would avoid building a scale for it. My rationale is both psychometric and political. Building scales for one sex occupations introduces problems of sex bias in scales and norms. A scale for "able-bodied seamen," based on a sample of 200 men in the occupation, is more likely to include or weight heavily a number of items which differentiate men from women than such a scale based on 200 women and 200 men. This problem might be solved if one had 400 men and 100 women by weighting the responses of each sex equally, but why build a scale for an occupation which is essentially a "male" occupation? The answer is to encourage more women to enter it. Analogously, we should build scales for occupations like child care worker to encourage more men to enter that occupation. This argument illustrates how the psychometric and political aspects of the problem interact in decision making. This phenomenon has, no doubt, been at work in interest measurement technology for years without anyone taking particular notice of it.

A general reference group for use in selecting items might be derived from the general population group used in developing homogeneous scales. Both Strong (1943) and Campbell (1971) have described problems in formulating a general reference group. Occupational level and culture as related to year of testing have been shown to effect the

reference group. Thus, the best reference group for selecting or weighting items on the occupational scales would become a subject for investigation.

Reliability

The scales developed must measure reliably over time if they are to be put into use. Appropriate reliability studies would be done since it has been demonstrated that interests can be measured reliably for both men and women (Campbell, 1966; 1969). I would anticipate no problems in this area.

Norms

I would provide general population norms for the homogeneous scales. At least high school, college age, and adult norms would be needed. The procedure used in item selection should minimize sex differences on these scales but whether or not they would eliminate them is a question which the data must answer. My guess is that I would still find women scoring [illegible] somewhat higher on the social or some corresponding scale and men scoring higher on mechanical or its corresponding scale.

Whatever my findings, I would build a periodic norming procedure into my organizational and fiscal structure. Too many test developers provide new norms only after it is clear that the old ones are outmoded or inadequat Realistically, this practice is understandable, but it contributes to poor testing practice. Ethically, it seems wiser to build the cost of up-to-date norms into the charge for the test, because it is an important service to the user.

For the occupational scales I would provide occupational norms. The evidence suggests less change over timewithin occupational groups than within total population so I would be less concerned about renorming these scales.

Validity

The occupational scales would presumably have concurrent validity because they were developed using occupational groups. Tilton's overlap (1937) would be computed between each occupational group and the general reference group to establish the concurrent validity of the occupational scales. Concurrent validity for the homogeneous scales, which would hopefully be related to Holland's types, could be established by scoring the occupational samples on them. The high scores of each occupational group would be expected to correspond to those listed for the occupation in Holland's Occupations Finder (1970).

A major concern would be whether the developed scales predict future occupational behavior. The only way to examine future behavior, assuming that the items selected for the scales had never been used together before, is to do a large scale longitudinal study. Ideally, the subjects originally studied would be of various ages, not just college freshmen. On followup, the predictive power of both types of scales, homogeneous and occupational, would be assessed by age at first testing and by sex. The basic data might take the form suggested by Figure 2. Hit rates or proportions of correct predictions would be calculated. Two sets of decision rules would be necessary, one for determining how individuals would be placed in each predictive category, and one for determining what kind of vocational behavior would be declared a correct prediction. Individuals could be placed in predictive categories on the basis of their highest scores, on the basis of scoring high on a scale compared to the general population, or on the basis of scoring high on a scale compared to people in a relevant occupation. Actually, calculating hit rates using various sets of decision rules

for making predictions would also provide valuable information about the various interpretive strategies.

Decision rules for assessing hits based on homogeneous scales might be based on predictions derived from either Holland's Occupations Finder or Cole's structural model as criteria, although a failure to find a high proportion of hits in this case might be interpreted as a lack of validity in either the predictor scales, the criteria, or both. It appears to be much easier to assign a hit or correct prediction when occupational scales are used because it is easy to tell whether the individual is in the occupation named by the scale or not. However, hits have usually been attributed to occupational scales when an individual entered the occupation named by the scale or a closely related occupation (Campbell, 1966 (b); Harmon, 1969).

One problem which must be addressed in assessing the predictive validity of interest inventories for women is that not all women work. The number and proportion of working women is increasing (U.S. Department of Labor, 1969) but Kriger's recent research (1972) suggested that the basic career decision made by women is whether or not to work. Perhaps my interest inventory should have a scale to predict orientation toward career commitment or away from it for both women and men. Schissel (1968) was able to build such a scale for women using the Strong Vocational Interest Blank for men. If a career orientation scale were available, non-working could be a predicted criterion. However, until the care of young children is shared equally by males and females in our society, there will probably be many young women with high career aspirations at home caring for children.

Harmon (1969) employed a strategy for assessing predictive validity which removed non-career women from consideration and gave career women who were temporarily out of the job market credit for their career commitment. In the proposed follow-up, the same latitude could be extended in assigning men to criterion categories.

A related problem is that many women do, unfortunately, enter occupations out of convenience or need. A large proportion of families with incomes below the poverty level are headed by women (United States Department of Labor, 1969). Few of them, left with families, can pay much attention to their interests and to long range planning. Circumstances like these also befall men but they seem more common for women. If they are, in fact, more common for women, we would expect a longitudinal follow-up of the predictive accuracy of an interest inventory to show better prediction for men than women. Thus, if society offers a wider range of choices to one sex or the other, we might expect better prediction of job criteria from interests for that sex, unless of course, there is an interaction between the opportunities society offers and the development of interests. Both Roe (1956) and Holland (1973) suggest that there is such an interaction. To explore cultural change and its interaction with interests our follow-up studies, as well as norming efforts, would have to be put on a schedule. Actually, they might go hand in hand, with new norm groups tested every five years and followed up 10, 15 and 20 years later.

Interpretation

Remembering that this fantasy is predicated on unlimited funds, I would do some unusual things to insure that the interest inventory I had developed would be well used. I would substantially

discount the cost of materials and scoring to every qualified user who attended regular workshops on its use and/or passed periodic examinations which showed that they understood vocational interest technology in general, and new developments in particular. Goldman (1972,a; 1972, b) has so despaired of the counselor's competency in using tests that he has advocated that counselors stop using them. I would attempt to educate counselors, instead.

Interpretative materials available for clients would stress the interaction between culture and testing. They would point out the necessity for renorming and the possibility that people in general might change over the next five years. For homogeneous and occupational scales, they would clearly present what validity data is now available and what is needed in the future to be more certain about the predictions we make. They would give the client enough information about what is and is not known, to make a decision as to how much weight to place on the scores. This procedure would be difficult to implement but it puts the responsibility where it belongs, on the client. The counselor whose needs are met by knowing more than the client and grandly unraveling the mystery before the clients' eyes, would not be very comfortable with either the ambiguity or the appeal to the client as counselor which these materials would contain.

Epilogue to the Fantasy

All this ambiguity makes me quite uncomfortable, too. None of the elements of my fantasy are new or terribly creative, but I am impatient because the answers are not yet available. We do not know the effect of balancing items in an interest item pool according to sexual response item differences. We do not know the effect of developing occupational scales with criterion groups of men and women, equally

in a structural model has predictive validity. We do not know much about how interests develop and change. While the ambiguities make me intellectually uncomfortable, I am even more uncomfortable with my clients.

None of the measurement procedures available to me provide me with both a lack of sexual bias and good predictive validity. Apparently, I will have to live out my own fantasy and be absolutely honest with my clients about the sexual bias in items, scales, and norms in the interest inventories I ask them to complete, and about the uncertain validities associated with using "wrong" sex scales or with using Cole's structural model.

When a fantasy is shared, what is the expected result? Approval, argument, or attention? This author will settle for action. I am not so concerned that we settle the questions I have raised in the direction I have predicted, but that we investigate whether our view of the world of work as dichotomized by sex is really necessary in interest measurement. This basic question must be answered before we know how to define sex bias in interest measurement more clearly and eliminate it more surely. To that end the following guidelines are dedicated.

Guidelines

In summary, the following guidelines can serve as both a means of assessing current interest inventories for sex bias as well as a plan for action for developing new interest inventories which are free of bias.

1. The content of interest inventory items should not imply that any occupation or activity is more appropriate for one sex than the other.
2. The pool of interest inventory items which make up an interest inventory should be appropriate for both sexes and used for both sexes.
3. Because there are sex differences in item responses, the item pool should contain equal numbers of items which are more highly endorsed by men than women and items which are more highly endorsed by women than men.
4. Groups used in developing scales should be composed of men and women in equal proportions or the effects of both groups should be statistically equalized.
5. Scores on the same scales should be available for both men and women (not just scores on scales which have the same names but which use different items).
6. Norms on homogeneous interest scales should not be presented by age or sex without evidence about the predictive value of a "high" score as related to such norm groups.
7. Norm groups for occupational scales should contain men and women in equal proportions or the effects of both groups should be statistically equalized.
8. Publishers of interest inventories should have a preannounced plan of periodically renorming homogeneous scales.
9. Publishers of interest inventories should have a preannounced plan for periodic studies of the predictive validities of both homogeneous and occupational interest scales.

10. Published interpretive materials should indicate clearly that sex norms for homogeneous scales do not imply anything about the predictive validity of the scales and encourage clients to evaluate their scores in relationship to both sexes.
11. Published interpretive materials should indicate clearly that using occupational scales developed for one sex with the other sex is of questionable validity because the underlying technology tends to maximize sex differences unduly.
12. Publishers of interest inventories which do not meet these guidelines should show evidence that attempts are being made to define, study, and eliminate sources of sex bias in their instruments.

References

American College Testing Program. Handbook for the ACT Career Planning Program. Iowa City, Iowa; ACT, 1972.

Association for Measurement and Evaluation in Guidance Commission on Sex Bias in Measurement. AMEG Commission report on sex bias in interest measurement. Measurement and Evaluation in Guidance, 1973, 6, 171-177.

Astin, H. S. & Myint, T. Career development of young women during the post high school years. Journal of Counseling Psychology, 1971, 18, 369-393.

Bird, C. Born Female. New York: David McKay, 1968.

Buros, O. K. (ed.). The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.

Buros, O. K. (ed.). The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.

Campbell, D. P. Strong Vocational Interest Blanks Manual (revised from E. K. Strong, Jr.) Stanford: Stanford University Press, 1966. (a)

Campbell, D. P. Occupations ten years later of high school seniors with high scores on the SVIB life insurance salesman scale. Journal of Applied Psychology, 1966, 50, 369-372. (b)

Campbell, D. P. Changing Patterns of Interests Within the American Society. Measurement and Evaluation in Guidance, 1968, 1, 36-49.

Campbell, D. P. 1969 Supplement to the Strong Vocational Interest Blanks Manual. Stanford, California: Stanford University Press, 1969.

Campbell, D. P. Handbook for the Strong Vocational Interest Blank. Stanford, California: Stanford University Press, 1971.

Campbell, D. P. Manual for the Strong Campbell Interest Inventory. Stanford, California: Stanford University Press, in press.

Campbell, D. P., & Holland, J. L. A merger in vocational interest research: Applying Holland's theory to Strong's data. Journal of Vocational Behavior, 1972, 2, 353-376.

Cole, N. S. On measuring the vocational interests of women. Journal of Counseling Psychology, 1973, 20, 105-112.

Cole, N. S. & Cole, J. W. An Analysis of spatial configuration and its application to research in higher education. ACT Research Report No. 35. Iowa City, Iowa: American College Testing Program, 1970.

Cole, N. S. & Hanson, G. R. An analysis of the structure of vocational interests. Journal of Counseling Psychology, 1971, 18, 478-486.

Cole, N. S., Whitney, D. R., & Holland, J. L. A spatial configuration of occupations. Journal of Vocational Behavior, 1971, 1, 1-9.

Contse, F. L. The effects of sexual role-playing on responses to the Strong Campbell Interest Inventory. Unpublished Masters Thesis: University of Wisconsin-Milwaukee, 1974.

Crissey, W. J. E. & Daniel, W. J. Vocational Interest Factors in Women. Journal of Applied Psychology, 1939, 23, 488-494.

Darley, J. G. & Hagenah, T. Vocational Interest Measurement. Minneapolis: University of Minnesota Press, 1955.

D'Costa, A. G. OVIS-A non-predicting device. Measurement and Evaluation in Guidance, 1972, 5, 411-419.

Diamond, E. E. Occupational interests: male-female or high level - low level dichotomy? Journal of Vocational Behavior, 1971, 1, 305-315.

Dolliver, R. H. Likes, dislikes, and SVIB scoring. Measurement and Evaluation in Guidance, 1968, 1, 73-80.

Firestone, S. The dialectic of sex: The case for feminist revolution. New York: William Morrow and Co., 1970.

Goldman, L. Tests and counseling: the marriage that failed. Measurement and Evaluation in Guidance , 1972, 4, 313-220. (a)

Goldman, L. It's time to put up or shut up. Measurement and Evaluation in Guidance, 1972, 5, 420-423, (b).

Hansen, J. C., and Johansson, C. B. The application of Holland's Vocational Model to the Strong Vocational Interest Blank for Women. Journal of Vocational Behavior, 1972, 2, 479-493.

Harmon, L. W. Optimum criterion group size in interest measurement. Measurement and Evaluation in Guidance , 1968, 1, 65-72.

Harmon, L. W. The predictive power over 10 years of measured social service and scientific interests among college women. Journal of Applied Psychology, 1969, 53, 193-198.

Harmon, L. W. The anatomy of career commitment in women. Journal of Counseling Psychology, 1970, 17, 77-80.

Harmon, L. W. Sexual bias in interest measurement. Measurement and Evaluation in Guidance, 1973, 6, 496-501.

Hill, D. S. & Campbell, D. P. The changing interests of women. Unpublished paper read at Midwestern Psychological Association Convention in Chicago, Illinois, May, 1969.

Holland, J. L. The Psychology of Vocational Choice. Waltham, Mass.: Blaisdell , 1966.

Holland, J. L. Manual for the Vocational Preference Inventory. Palo Alt California: Consulting Psychologists Press, 1965a.

Holland, J. L. The Self-Directed Search. Palo Alto, California: Consulting Psychologists Press, 1965b.

Holland, J. L. Making Vocational Choices: A Theory of Careers. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Holland, J. L., Whitney, D. R., Cole, N. S., & Rochards, J. M., Jr. An empirical occupational classification derived from a theory of personality and intended for practice and research. ACT Research Report No. 29. Iowa City, Iowa: The American College Testing Program, 1969.

Johansson, C. B. & Harmon, L. W. Strong Vocational Interest Blank: One Form or Two? Journal of Counseling Psychology, 1972, 19, 404-41

Johnson, R. W. The use of interest inventories with women. University of Wisconsin Counseling Center Reports, 1970, 3, No. 13.

King, P, Norell, G., & Powers, P. Relationships between twin scales on the SVIB and the Kuder. Journal of Counseling Psychology, 1963, 10, 395-401.

Kriger, S. F. Need achievement and perceived parental child-rearing attitudes of career women and homemakers. Journal of Vocational Behavior, 1972, 2, 419-432.

Kuder, F. Kuder Occupational Interest Survey General Manual. Chicago Science Research Associates, 1970 (b).

Kuder, F. A note on the comparability of occupational scores from different interest inventories. Measurement and Evaluation in Guidance, 1969, 2, 94-100.

Kuder, F. Some principles of interest measurement. Educational and Psychological Measurement, 1970, 30, 205-226.

Lewis, E. C. Developing Women's Potential. Ames, Iowa: Iowa State University Press, 1968.

Millet, Kate. Sexual Politics. New York: Doubleday, 1969.

O'Shea, A. J., Lynch, M. D., & Harrington, T. F. A reply to Kuder's criticism of SVIB-KOIS comparative studies. Measurement and In Guidance, 1972, 5, 306-309.

O'Shea, A. J. & Harrington, T. F. Using the Strong Vocational Interest Blank and the Kuder Occupational Interest Survey, Form DD, with the same clients. Journal of Counseling Psychology, 1971, 18, 44-50.

Parsons, T. Social Structure and Personality. London: The Free Press, Collier-Macmillan, 1965.

Roe, A. The Psychology of Occupations. New York: John Wiley, 1956.

Schissel, R. F. Development of a career orientation scale for women. Journal of Counseling Psychology, 1968, 15, 257-262.

Schlossberg, N. K. & Goodman, J. Imperative for change: Counselor use of the Strong Vocational Interest Blanks. Impact, 1972, 2, 25-29.

Stanfiel, J. D. Administration of the SVIB Men's form to women counselees. Vocational Guidance Quarterly, 1970, 19, 22-27.

Strong, E. K., Jr. Vocational Interests of Men and Women. Stanford: Stanford University Press, 1943.

Strong, E. J., Jr. Vocational interests 18 years after college. Minneapolis: University of Minnesota Press, 1955.

Strong, E. K., Jr. Good and poor interest items. Journal of Applied Psychology, 1962, 46, 269-275.

Super, D. E. and Crites, J. O. Appraising Vocational Fitness. New York: Harper and Row, 1962.

Thurstone, L. L. A multiple factor analysis of vocational interests. Personnel Journal, 1931, 10, 198-205.

Tilton, J. W. Measurement of overlapping. Journal of Educational Psychology, 1937, 28, 656-662.

United States Department of Labor. 1969 Handbook on Women Workers. Women's Bureau Bulletin No. 294. Washington, U. S. Government Printing Office, 1969.

Women on Words and Images. Dick and Jane as Victims. Princeton, New Jersey: Women on Words and Images, 1972.

Zytowski, D. Relationships of equivalent scales on three interest inventories. Personnel and Guidance Journal, 1968, 47, 44-49.

Zytowski, D. G. A test of criterion group sampling errors in two comparable interest inventories. Measurement and Evaluation in Guidance, 1969, 2, 37-40.

Zytowski, D. G. A concurrent test for accuracy of classification for the Strong Vocational Interest and Kuder Occupational Interest Survey. Journal of Vocational Behavior, 1972, 2, 245-250.

Zytowski, D. G. The Kuder Occupational Interest Survey. In Zytowski, D. G. (ed.) Contemporary Approaches to Interest Measurement. Minneapolis: University of Minnesota Press, 1973.

Zytowski, D. G. The predictive validity of the Kuder Preference Record over a 25 year span. Measurement and Evaluation in Guidance. 1974, 7, in press.

Table 1

Items Appearing on Three Tentative SCII Scales
For Veterinarians (Items 1-131, Occupations)

Items	Direction "Like" is Scored	Scales on Which Items Appear F (23%)	M (18%)	F&M (15%)
2 Advertising Executive	-		X	X
10 Auctioneer	+		X	X
13 Author of technical books	+	X		X
19 Biologist	+	X	X	X
29 Church worker	+	X		
30 City or State Employee	-	X	X	X
31 City Planner	-		X	X
34 Computer Operator	-		X	
41 Dentist	+		X	X
42 Designer, Electronic equipment	-		X	
43 Dietician	-		X	
46 Editor	-		X	
49 Elementary School Teacher	-	X		X
50 Employment Manager	-	X	X	X
52 Farmer	+	X	X	X
59 High School Teacher	-	X		
60 Home Economics Teacher	-	X		X
70 Labor Arbitrator	-		X	
71 Laboratory Technician	+	X		X
78 Manager Child Care Center	-	X		X
79 Manager Women's Style Shop	-	X		X
85 Newspaper Reporter	-		X	X
86 Nurse	+		X	X
88 Office Clerk	-	X		
89 Office Manager	-			X
92 Pharmacist	+		X	X
94 Physician	+	X	X	X
95 Playground Director	-	X		
98 Politician	-		X	X

Items		F	M	F&M
99 Private Secretary	-	X		
104 Public Relations Director	-	X		X
105 Rancher	+	X	X	X
107 Receptionist	-	X		
109 Sales Manager	-		X	X
111 Scientific Illustrator	+	X		X
112 Scientific Research Worker	+	X		X
115 Social Worker	-	X	X	X
121 Surgeon	+	X	X	X
127 Vocational Councelor	-	X		X
130 X-Ray Technician	+		X	X
131 YMCA/YWCA Staff Member	-	X		
		25	23	29

Figure 1

Responses to the Item "Biologist"

	Like %	Indifferent %	Dislike %
Veterinarians (Female)	90	9	1
Women-in-General	46	29	25
Differences	+ 44	- 20	- 24
Scale Weights	+ 1	- 1	- 1

Figure 2

Predictive Validity of Interest Scales
Established by Hit Rates

Predictions (by scale)	Number of Predictions by Age Group At Testing	N	Percentage of Hits: Sex: Male	Female	Total
Social (homogeneous)	High School				
	College				
	Adult				
	Total				
Social (Occupational)	High School				
	College				
	Adult				
	Total				
Investigative (homogeneous)	High School				
	College				
	Adult				
	Total				
Scientific (Occupational)	High School				
	College				
	Adult				
	Total				

etc.